
Ouverture des données de recherche

Guide d'analyse du cadre juridique en France*



[Contenu sous licence Creative Commons Attribution 4.0 International \(CC BY 4.0\).](https://creativecommons.org/licenses/by/4.0/)

* Le présent guide est issu des réflexions d'un groupe de travail inter-organismes. Il ne prétend pas à l'exhaustivité et est fourni uniquement à titre d'information. Il ne saurait en tout état de cause se substituer aux politiques d'établissements, au respect des dispositions législatives ou réglementaires et au respect de la jurisprudence applicable en la matière. Ce guide peut évoluer.

Membres du groupe de travail : BECARD Nicolas (INRA), CASTETS-RENARD Céline (UT1), CHASSANG Gauthier (Inserm, Membre de la Plateforme Genotoul Societal), COURTOIS Mary-Audrey (INRA), DANTANT Martin, GANDON Nathalie (co-animatrice, INRA), MARTIN Caroline (Irstea), MARTELLETTI Andrea (stagiaire INRA, M2 droit et Informatique), MENDOZA-CAMINADE Alexandra (UT1), MORCRETTE Nathalie (co-animatrice, INRA), NEIRAC Claire (Cirad) - *Guide retravaillé par MARTELLETTI Andrea dans le cadre de son stage de fin d'études de M2 Droit et informatique.*

Préface

Le contexte technologique dans lequel s'inscrivent les activités de recherche est en plein bouleversement : les moyens d'investigation (collecte, analyse, traitement) font massivement appel aux technologies numériques et les données générées réclament attention.

Mais, pour faire face aux enjeux de société qu'elle interroge, la science se doit d'être également plus interdisciplinaire, plus collaborative, plus participative, plus transparente, plus ouverte et pluscirculante !

A l'instar des publications, qui voient aujourd'hui leur ouverture consacrée par le mouvement Open Access, les données de recherche apparaissent comme un patrimoine essentiel de connaissances à partager et réexploiter à des fins de recherche, d'enseignement ou de vulgarisation scientifique.

Les politiques nationales, européennes et internationales convergent aujourd'hui vers un soutien en faveur du libre accès aux productions scientifiques, et les infrastructures nécessaires à cet objectif se mettent en place.

Construire de bonnes pratiques partagées est essentiel au fonctionnement harmonieux des groupes sociaux et les communautés disciplinaires de chercheurs n'échappent pas à cette règle. Ce mouvement vertueux sera d'autant facilité qu'il se construira dans un paysage juridique éclairci, adapté aux nécessités de la science moderne.

Il convenait de toiletter l'arsenal juridique à la disposition des chercheurs pour satisfaire aux nécessités de la recherche du 21^{ème} siècle, consacrant le partage et la réutilisation des productions scientifiques en principe de base, pour une participation de la science à une société plus ouverte, innovante et inclusive.

C'est chose faite avec la loi française « pour une république numérique » et c'est également l'ambition portée par le projet de révision de la directive européenne « droits d'auteur et droits voisins ».

Les activités de recherche, en prise directe avec les besoins de la société (transport, vieillissement, énergie, alimentation, sécurité,...) méritent de s'inscrire dans un droit positif et ne peuvent dorénavant se contenter d'un régime d'exceptions dans une législation de propriété intellectuelle centrée sur la protection de la création.

Sans être (pour l'instant) qualifiées de « biens communs », les données de recherche gagnent dans la loi « pour une république numérique » un statut d'objets partageables et réutilisables dont la circulation est facilitée. Il est temps de se saisir de cette opportunité tout en anticipant ses limites et contraintes.

Ainsi, la nature de certaines données et le contexte partenarial et concurrentiel dans lequel s'inscrivent aujourd'hui les activités de recherche créent une tension entre cette volonté d'ouverture et les contraintes qui s'imposent par ailleurs aux établissements de recherche, notamment les objectifs de valorisation des résultats de la recherche. C'est pourquoi il convient de porter toute l'attention nécessaire aux catégories de données qui doivent désormais être ouvertes, certaines ne pouvant pas faire l'objet d'une diffusion publique sans précautions préalables. Fruit d'un effort coordonné de plusieurs juristes de l'enseignement supérieur et de la recherche français, le présent document vise à donner aux chercheurs les éléments clés qu'il faut prendre en compte pour décider d'une ouverture des données de la recherche, quel que soit le type de recherche mise en œuvre, de façon à respecter les nouvelles dispositions juridiques dans le domaine.

L'objectif principal de ce guide étant la vulgarisation, il atteindra son but s'il suscite interrogations, débats, commentaires, montrant ainsi l'intérêt des chercheurs pour la chose juridique et son application dans leurs activités professionnelles.

Les membres de BSN10 saluent le travail coopératif accompli dans la production de ce guide, et recommandent vivement son partage et sa consommation sans modération !

Les membres de BSN10



Décembre 2016

SOMMAIRE :

PREAMBULE	3
QU'EST-CE QUE L'OPEN DATA ?	3
NOTION D'OPEN DATA	3
OPEN DATA DES DONNEES DE LA RECHERCHE	6
1-LES DONNEES QUE J'EXPLOITE SONT-ELLES CONCERNEES PAR L'OPEN DATA ?	7
1.1 COMMUNICATION OBLIGATOIRE.....	7
1.1.1 DONNEES GEOGRAPHIQUES :	7
1.1.2 DONNEES ENVIRONNEMENTALES RELATIVES A DES EMISSIONS DE SUBSTANCES DANS L'ENVIRONNEMENT	8
1.2 COMMUNICATION INTERDITE PAR PRINCIPE	9
1.2.1 DONNEES PRESENTANT DES RISQUES POUR LA PROTECTION DU SECRET DE LA DEFENSE NATIONALE :	9
1.2.2 DONNEES PRESENTANT DES RISQUES POUR LA SECURITE PUBLIQUE DE L'ETAT OU POUR LA SECURITE DE L'ETABLISSEMENT	9
1.2.3. SECRETS PROFESSIONNELS.....	10
1.3 COMMUNICATION SOUS CONDITIONS	13
1.3.1 DONNEES PRESENTANT DES RISQUES POUR LA PROTECTION DU POTENTIEL SCIENTIFIQUE ET TECHNIQUE DE LA NATION	13
1.3.2 LE CAS DES ZONES A REGIME RESTRICTIF (ZRR)	13
1.3.3. DONNEES PROTEGEES PAR LE DROIT D'AUTEUR ET AUTRES DROITS DE PROPRIETE INTELLECTUELLE	14
1.3.3.1 DROIT D'AUTEUR	14
1.3.3.2 AUTRES DROITS DE PROPRIETE INTELLECTUELLE	16
1.3.4 DONNEES RELATIVES AUX PERSONNES, A LA VIE PRIVEE.....	18
1.3.5 DONNEES STATISTIQUES.....	19
1.3.6 DONNEES LIEES A UN CONTRAT AVEC UN TIERS NON SOUMIS A UNE OBLIGATION DE SERVICE PUBLIC	19
1.4 CONCLUSION.....	20
2. COMMENT DIFFUSER LES DONNEES ?	21
2.1. LES GRANDS PRINCIPES :	21
2.1.1 ASPECTS TECHNIQUES DE L'OPEN DATA.....	21
2.1.2 ASPECTS JURIDIQUES : CHOIX D'UNE LICENCE DE DIFFUSION	22
2.2. LES DIFFERENTES MODALITES DE DIFFUSION :	23
2.2.1. LES PRATIQUES PAR DISCIPLINE	23
2.2.2. LA POLITIQUE D'ETABLISSEMENT	24
2.2.3. LE REPERTOIRE DES DONNEES PUBLIQUES	24
2.2.4. LA DEMANDE D'ACCES FORMULEE PAR UN TIERS	24
LOGIGRAMME DE COMMUNICABILITE D'UNE DONNEE	26
FICHE 1 : LES BASES DE DONNEES EN BREF	28
FICHE 2 : LES DONNEES PERSONNELLES	32
FICHE 3 : LES DONNEES STATISTIQUES	36
FICHE 4 : CONVENTION D'AARHUS SUR L'INFORMATION EN MATIERE D'ENVIRONNEMENT	38
FICHE 5 : TABLEAU COMPARATIF DES LICENCES CC4.0 ET CC0.....	40
FICHE 6 : TABLEAU COMPARATIF DES LICENCES GRATUITES ODBL ET ETALABAB	43
FICHE 7 : MISE EN PLACE D'UNE LICENCE PAR LA VOIE DU CONTRAT ELECTRONIQUE	45

Préambule

« La recherche publique a pour objectifs le développement et le progrès de la recherche dans tous les domaines de la connaissance ; la valorisation des résultats de la recherche ; le partage et la diffusion des connaissances scientifiques ; le développement d'une capacité d'expertise ; la formation à la recherche et par la recherche et l'organisation de l'accès libre aux données de la recherche »¹.

Ce guide sur l'ouverture (ou Open Data) des données de recherche a pour vocation d'accompagner les agents des établissements concernés (établissements d'enseignement et organismes de recherche) dans une démarche d'ouverture raisonnée des données de recherche en tentant de répondre aux questions les plus courantes auxquelles ils pourront être confrontés, que cette démarche soit volontaire et réponde aux objectifs de l'établissement ou qu'elle soit imposée par la réglementation. Il est précisé que le cadre légal est cité, lorsqu'il existe. L'attention du lecteur est toutefois attirée sur le paysage très mouvant du droit² sur ce sujet et sur la nécessité de se référer à la politique de son établissement en matière d'Open Data.

Qu'est-ce que l'Open Data ?

Notion d'Open Data

De façon générale, l'expression « Open Data » peut être définie comme une démarche de communication des documents ou données publics, afin qu'ils soient diffusés de manière structurée selon une méthode garantissant leur libre accès et leur réutilisation par tous, sans restriction technique, juridique ou financière injustifiée¹.

L'Open Data est régi en Europe par une Directive 2013/37/UE dite « PSI » pour « Public Sector Information » et, en France, par la loi dite « CADA³ » n°78-753 du 17 juillet 1978 (modifiée à plusieurs reprises) codifiée depuis le 19 mars 2016 dans le Code des relations entre le public et l'administration (articles L300-1 et suivants). Ces textes définissent, d'une

¹ [Article L112-1 du Code de la recherche](#)

² Ce guide fera l'objet d'une v2 en 2017 pour intégrer de façon plus complète la loi pour une République numérique n°2016-1321 du 7 octobre 2016 (notamment les décrets d'application de cette loi)

³ Commission d'Accès aux Documents Administratifs

part, un droit d'accès individuel aux documents administratifs pour le citoyen et, d'autre part, un droit de réutilisation de certains de ces « documents administratifs » pour tous.

Les « documents administratifs » visés par la loi française⁴ sont tous les documents quels que soient leur date, leur lieu de conservation, leur forme et leur support qui sont produits ou reçus, dans le cadre de leur mission de service public, par l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission.

La loi donne les exemples suivants : dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions, codes sources⁵ et décisions. Cette liste n'est pas limitative.

Pour la qualification des données émanant des établissements de recherche et enseignement en « documents administratifs », il y a lieu de retenir deux composantes :

1. Les données produites par les établissements de recherche et d'enseignement **dans le cadre de leur mission de service public** sont considérées comme des documents administratifs et sont donc communicables à toute personne qui en fait la demande, sauf exceptions légales.
2. Lorsque la loi parle de « documents administratifs », cela englobe également les **données produites par ces établissements, quelles qu'elles soient** : données brutes⁶, données élaborées et métadonnées. Cependant, la loi précise que ne peuvent être accessibles au public que les documents « achevés ». Par conséquent, tous les documents préparatoires ne sont pas considérés comme des documents administratifs. On peut donc en déduire que les cahiers de laboratoire sont exclus de la réglementation sur l'ouverture des données (cf. également infra la définition OCDE de données de la recherche).

Ce guide n'aborde pas tous les types de données mais uniquement celles que les équipes de recherche produisent le plus couramment. Certaines données (données de transport, données essentielles des marchés publics, etc.) qui font l'objet d'une obligation de diffusion ne sont pas abordées ici.

⁴ Article L300-2 du Code des relations entre le public et l'administration

⁵ Voir en ce sens le [jugement du tribunal administratif de Paris 10/03/2016](#) sur le code source du logiciel de calcul des impôts

⁶ Les données préliminaires sont des données préparatoires, préalables, nécessaires à la mise en place d'une expérimentation, d'un procédé, d'une enquête, etc : il ne s'agit pas de données de recherche. Les données brutes sont des données issues d'une l'expérimentation, d'un procédé, d'une enquête, etc : il peut s'agir de données de recherche communicables.

Quelles sont les conséquences de la qualification des données en « document administratif » ?

Il convient d'être particulièrement prudent lorsqu'il y a une publication scientifique et que l'éditeur impose le dépôt des données dans un entrepôt spécifique. Si c'est bien le scientifique qui décide du contenu de la publication, c'est, en revanche, à l'établissement de décider quelles données seront ouvertes, où elles doivent être déposées et sous quelles conditions.

Les décisions d'ouverture des données se prennent au niveau de l'établissement et non pas au niveau de l'agent.

1. Le principe posé par la réglementation –tant française qu'européenne- est celui de la **liberté d'accès** aux documents administratifs pour toute personne qui en fait la demande. Si les données ne font pas déjà l'objet d'une diffusion publique, tout un chacun (citoyen, association, entreprise, administration, etc.), peu importe sa nationalité, peut demander à l'établissement/l'organisme qui a produit les données la simple consultation de celles-ci en vue d'une utilisation privée ou interne (c'est-à-dire non communiquée au public, à usage personnel ou interne). Il y a cependant des exceptions pour certains types de données, d'où la nécessité de procéder à l'analyse de la nature des données. Les différentes étapes qui sont décrites dans ce guide fournissent une méthode pour effectuer cette analyse.

2. Pour certaines données, en sus du droit d'accès aux documents, il existe un droit de **réutilisation**⁷. L'instauration de ce droit à la réutilisation d'informations publiques témoigne de la volonté des institutions européennes et françaises d'amplifier le mouvement d'Open Data, dont l'objet est l'accès libre aux informations du secteur public au bénéfice du citoyen en vue de :

- garantir la transparence de l'Etat ;
- valoriser les données publiques ;
- favoriser le développement d'activités privées et l'émergence de nouveaux opérateurs économiques.

Parmi les données accessibles, certaines peuvent être réutilisées par tout acteur qui souhaite en faire un usage différent que celui répondant à la mission de service public initiale. Aucune restriction d'usage ne peut alors être apportée par l'Administration. Cependant, toutes les données accessibles ne sont pas forcément réutilisables (cf. la méthode d'analyse ci-après).

A ce sujet, depuis la loi pour une République Numérique du 7 octobre 2016, il est prévu que la réutilisation de données issue d'une activité de recherche est libre si :

- ces données sont issues d'une activité de recherche financée au moins pour moitié par des fonds publics ; et
- ces données ne sont pas protégées par un droit spécifique ;

⁷ Art L321-1 du code des relations entre le public et l'administration

- et ces données ont été rendues publiques par le chercheur ou l'établissement.

La distinction entre droit d'accès et droit de réutilisation est très importante. En effet, une personne qui a droit d'accès à des données n'a pas nécessairement le droit de les réutiliser.

Open Data des données de la recherche

Les droits et principes définis ci-dessus en matière d'open data s'appliquent pleinement aux données de la recherche.

En matière de données de la recherche, il existe une définition (qui n'a pas de valeur légale) donnée par l'OCDE en 2007 dans un document relatif aux principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics. Selon l'OCDE, les « *données de la recherche* » sont définies comme « *des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche. Un ensemble de données de recherche constitue une représentation systématique et partielle du sujet faisant l'objet de la recherche. Ce terme ne s'applique pas aux éléments suivants : carnets de laboratoire, analyses préliminaires et projets de documents scientifiques, programmes de travaux futurs, examens par les pairs, communications personnelles avec des collègues et objets matériels (par exemple, les échantillons de laboratoire, les souches bactériennes et les animaux de laboratoire tels que les souris).* »

En conclusion, les deux conditions préliminaires pour diffuser mes données selon les principes de l'Open Data sont :

- des **données achevées** ;
- des **données réalisées dans le cadre de la mission de service public** de mon établissement (ceci est particulièrement important pour les EPIC).

MAIS d'autres conditions doivent aussi être réunies, c'est pourquoi je dois vérifier la nature de mes données selon les étapes ci-dessous.

1-Les données que j'exploite sont-elles concernées par l'Open Data ?

Attention à ne pas confondre la communication obligatoire de jeux de données et le droit d'alerte institué par la loi n°2013-316 du 16 avril 2013, qui dit que : « Toute personne physique ou morale a le droit de rendre publique ou de diffuser de bonne foi une information concernant un fait, une donnée ou une action, dès lors que la méconnaissance de ce fait, de cette donnée ou de cette action lui paraît faire peser un risque grave sur la santé publique ou sur l'environnement. L'information qu'elle rend publique ou diffuse doit s'abstenir de toute imputation diffamatoire ou injurieuse. ». Une telle diffusion d'information ne constitue pas une démarche d'Open Data puisque l'information s'adresse avant tout aux autorités publiques et que sa diffusion n'entraîne pas automatiquement un droit à réutilisation.

Pour en savoir plus: Guide « La Directive Inspire pour les néophytes », page 10 Site internet naturefrance

Les données produites par les établissements de recherche et d'enseignement dans le cadre de leur mission de service public sont considérées comme des documents administratifs et sont donc **communicables à toute personne qui en fait la demande, sauf exceptions légales**.

Cette communication peut être obligatoire (1.1), interdite (1.2) ou soumise à certaines conditions (1.3) : cela dépend de la nature de mes données.

1.1 Communication Obligatoire

Certaines données géographiques et environnementales, doivent obligatoirement être ouvertes au public.

1.1.1 Données géographiques :

Pour les données géographiques **qui sont disponibles sous format électronique**, c'est la directive européenne INSPIRE⁸ qui s'applique. Elle impose aux autorités publiques :

- de rendre accessibles au public leurs données environnementales géographiques en publiant sur internet ces données et les métadonnées correspondantes ;
- de partager leurs données environnementales géographiques entre elles.

Ces données géographiques se décomposent en trois groupes principaux :

- Les données nécessaires au repérage sur le territoire. *Exemple : hydrographie.*
- Les données générales complémentaires. *Exemples : altimétrie, géologie.*
- Les données thématiques telles que bâtiments, vocation des sols, santé et sécurité des personnes, services d'utilité publique et services publics, etc.

⁸ [Directive INSPIRE 2007/2/CE du Parlement européen et du Conseil du 14 mars 2007.](#)

1.1.2 Données environnementales relatives à des émissions de substances dans l'environnement

Concernant les données environnementales, le principe d'accès à ces données est fixé par la Convention d'Aarhus⁹, qui déclare que « toute personne a le droit d'être informée, de s'impliquer dans les décisions et d'exercer des recours en matière d'environnement ». Ce principe a été repris par la législation européenne¹⁰, puis par la législation nationale¹¹.

Pour en savoir plus :
[Fiche 5 : Convention Aarhus](#)

Parmi les données environnementales, le cas des informations relatives à des émissions de substances dans l'environnement¹² est particulier puisque ces informations peuvent être communiquées à celui qui en fait la demande même si elles sont préliminaires (= non achevées).

Ces données recouvrent les émissions de substances susceptibles d'avoir des incidences sur l'air, l'eau, le sol, etc.

*Exemples : enquête environnementale suite à une déclaration de mortalité massive d'abeilles¹³ => communication obligatoire
Enquête sur une épidémie de légionellose¹⁴ => communication refusée (car les données recueillies n'étaient pas relatives à l'environnement)*

En outre, les cas de refus de communication sont extrêmement limités. Le refus de communiquer les données n'est justifié que si :

- la conduite de la politique extérieure de la France, la sécurité publique ou la défense nationale sont en cause ;
- des procédures juridictionnelles peuvent être impactées ;
- la communication porte atteinte à des droits de propriété intellectuelle.

En conclusion :

Je possède des données géographiques informatisées

=> je dois les diffuser.

Je possède des données relatives à des émissions de substances dans l'environnement

=> je dois les communiquer à ceux qui les demandent, sauf cas exceptionnels.

⁹ La convention internationale d'Aarhus a été signée le 25 juin 1998

¹⁰ Directives européennes [2001/42/CE du 27 juin 2001](#), [2003/4/CE du 28 janvier 2003](#), [2003/35/CE du 26 mai 2003](#), [2000/60/CE du 23 octobre 2000](#)

¹¹ [Loi n° 2005-1319 du 26 octobre 2005](#), [décret n°2002-1187 du 12 septembre 2002](#), [décret n° 2006-578 du 22 mai 2006](#) codifiés dans le code de l'environnement

¹² [Article L124-5 Code de l'Environnement](#)

¹³ [Avis CADA n° 20130750 du 28/03/2013](#)

¹⁴ [Avis CADA n° 20090310 du 26/02/2009](#)

1.2 Communication interdite par principe

1.2.1 Données présentant des risques pour la protection du secret de la défense nationale :

Selon les dispositions du code pénal¹⁵, les procédés, objets, documents, données ou fichiers informatisés intéressant la défense nationale peuvent faire l'objet d'une classification au titre de la protection du secret de la défense nationale (confidentiel défense, secret défense). Cette classification concerne, au-delà du domaine militaire, tous les champs d'activités nécessitant de limiter l'accès à la connaissance d'un contenu, pour ne pas créer des préjudices graves pour la sécurité nationale. La diffusion de tout ou partie de ces informations sans autorisation de l'émetteur est considérée comme un délit.

A titre d'exemple, peuvent être classifiés au titre de la protection du secret de la défense nationale :

- Les documents relatifs aux plans gouvernementaux ;
- Les documents relatifs aux opérateurs d'importance vitale ;
- Les études financées par le ministère de la défense ou l'union européenne pour lesquelles il a été demandé une classification au titre de la défense.

1.2.2 Données présentant des risques pour la sécurité publique de l'Etat ou pour la sécurité de l'établissement

L'impératif de protection de la sécurité publique de l'Etat ou de la sécurité des biens et des personnes d'un établissement peut dans certains cas faire obstacle à la communication d'informations et données.

Les motifs de refus du droit d'accès, compte tenu de ces risques, sont :

- la sécurité publique de l'Etat.
Exemples : la liste des laboratoires ayant de l'anthrax ne doit pas être ouverte pour prévenir les attentats tels que celui des « enveloppes piégées » envoyées suite aux attentats du 11 septembre 2001 ; la liste des réservoirs d'eau potable ne doit pas être ouverte pour prévenir des attentats à la contamination, etc.
- la sécurité des biens de l'établissement

¹⁵ Article 413-9 Code pénal

Exemples : liste des laboratoires dont les recherches peuvent être soumises à contestation (OGM, expérimentation animale) pour prévenir des actions malveillantes ; plan complet des sites pouvant révéler les points de vulnérabilité pour prévenir des infractions; etc.

- **la sécurité des personnes.**

Exemples : diffusion sur internet d'informations permettant de fabriquer une bombe artisanale ; liste de personnels d'installations sensibles ; etc.

1.2.3. Secrets professionnels

[L'article L311-5 2° h du Code des relations entre le public et l'administration](#) indique que les documents administratifs dont la diffusion porterait atteinte aux autres secrets protégés par la loi ne sont pas communicables.

Avertissement :
La liste des secrets n'est pas exhaustive.
Par exemple, le secret des correspondances n'est pas traité par le guide.

Parmi ces secrets, le plus fréquemment rencontré par les établissements publics de recherche et d'enseignement supérieur est le secret professionnel.

Le secret professionnel est une notion large et générique qui comprend notamment:

- **Le secret médical :** cette obligation s'applique aux informations relatives à la santé, au comportement, à la situation familiale d'une personne physique, etc. => en cas de demande de communication de ce type de données, je ne peux pas faire droit à cette demande sans l'accord exprès de la personne concernée.
- **Le secret défense => (Cf. 1.2.1).**
- **Le secret des affaires.**

Le secret des affaires est une notion qui n'est pas clairement définie (notamment par le droit français) mais les éléments suivants sont à prendre en compte :

En Europe :

Les parlementaires européens viennent de s'accorder au travers d'une directive¹⁶ qui énonce 3 conditions à réunir pour qu'une information constitue un secret des affaires :

1. Les informations sont secrètes en ce sens que, dans leur globalité ou dans la configuration et l'assemblage exacts de leurs éléments, elles ne sont pas généralement connues de personnes appartenant aux milieux qui s'occupent normalement

¹⁶ [Directive 2016/943 du 8 juin 2016](#)

du genre d'informations en question, ou ne leur sont pas aisément accessibles;

2. Les informations ont une valeur commerciale parce qu'elles sont secrètes;
3. Les informations ont fait l'objet, de la part de la personne qui en a licitement le contrôle, de dispositions raisonnables, compte tenu des circonstances, destinées à les garder secrètes.

En France :

La CADA¹⁷ a tenté de définir les secrets industriels et commerciaux : il s'agit d'éléments sensibles ayant un impact sur l'environnement concurrentiel de l'établissement et de ses partenaires. Selon cette autorité : « *La communication des documents contenant des informations dont la divulgation pourrait porter atteinte au secret industriel et commercial est réservée aux seuls intéressés. La notion de secret industriel et commercial recouvre trois catégories de données* »¹⁸ :

- **le secret des procédés ;**
- **le secret des informations économiques et financières ;**
- **le secret des stratégies commerciales ou industrielles.**

La loi réprime par des sanctions pénales, le fait de divulguer des informations¹⁹ couvertes par « les secrets » cités ci-dessus.

Toutefois, le secret professionnel peut être levé sur autorisation de la personne concernée par l'information et sous réserve que soient préservés :

- La protection des personnes.
Exemple : révélation de maltraitances.
- La santé publique.
Exemple : révélation de maladies nécessitant une surveillance.
- L'ordre public et le bon déroulement des procédures de justice.
Exemples : dénonciation de crimes ou de délits, témoignages en justice.

Les données couvertes par le secret des affaires (y compris donc les secrets industriels et commerciaux) peuvent être des données développées par l'établissement public (procédé, secret), mais aussi des données appartenant à des partenaires industriels qui ont été transmises à un établissement public à l'occasion, par exemple, d'une collaboration.

Pour en savoir plus :

Documents CADA
- [Le secret en matière commerciale et industrielle](#)
- [Les documents couverts par le secret en matière commerciale et industrielle.](#)

¹⁷ Commission d'Accès aux Documents Administratifs

¹⁸ Cette définition a été reprise par la loi pour une République Numérique du 7 octobre 2016 ([article L311-6 Code des relations entre le public et l'administration](#))

¹⁹ « La révélation d'une information à caractère secret par une personne qui en est dépositaire soit par état ou par profession, soit en raison d'une fonction ou d'une mission temporaire, est punie d'un an d'emprisonnement et de 15 000 euros d'amende ». [Article 226-13 du code pénal.](#)

En conclusion :

Je dois vérifier que les informations que je souhaite diffuser ne contiennent pas d'éléments pouvant :

- 1. relever du secret défense**
- 2. présenter des risques pour la sécurité publique ou celle de l'établissement**
- 3. relever du secret professionnel (médical, affaires, ...)**

L'interdiction de diffusion est le principe.

Mais si je souhaite tout de même diffuser :

Dans les cas 1 et 2, je contacte le fonctionnaire de sécurité et/ou la Direction Générale de l'établissement.

Dans le cas 3, je contacte le service juridique.

1.3 Communication sous conditions

1.3.1 Données présentant des risques pour la protection du potentiel scientifique et technique de la nation

Le potentiel scientifique et technique est constitué de l'ensemble des biens matériels et immatériels propres à l'activité scientifique fondamentale et appliquée et au développement technologique. Les éléments essentiels du potentiel constituent des intérêts fondamentaux de la nation définis dans le code pénal²⁰.

Un dispositif réglementaire rénové en 2012 a pour objectif la protection de ce potentiel scientifique et technique de la nation (PPST). Il vise à protéger l'accès aux savoirs et aux savoir-faire et aux technologies des établissements privés ou publics localisés sur le territoire national, lorsque leur détournement ou leur captation pourraient :

- soit porter atteinte aux intérêts économiques de la nation,
- soit renforcer des arsenaux militaires étrangers ou affaiblir les capacités de défense de la France,
- soit contribuer à la prolifération des armes de destruction massive (nucléaire, chimique ou biologique) et de leurs vecteurs,
- soit de favoriser les actions malveillantes (terroristes) sur le territoire national ou à l'étranger.

Si mon laboratoire peut contribuer à l'un de ces risques, on parle d'unité protégée.

Il est nécessaire de vérifier que la diffusion des données ne porte préjudice pas aux potentiels scientifiques et techniques de mon unité, de mon établissement ou de la Nation.

Il est recommandé de demander l'avis d'un responsable de service ou de l'unité avant diffusion, qui peut interroger le fonctionnaire de sécurité ou d'autres services compétents.

1.3.2 Le cas des zones à régime restrictif (ZRR)

Si les savoirs et savoir-faire de mon laboratoire peuvent contribuer à l'un des principaux risques énoncés au 1.3.1 ci-dessus, le ministère de tutelle de mon établissement peut avoir identifié un besoin supplémentaire de

²⁰ [article 410-1 Code pénal](#)

protection et avoir demandé la création de zones à régime restrictif (ZRR²¹).

L'accès physique ou numérique à ces zones est soumis à autorisation. La diffusion de données issue de ces zones doit au préalable avoir été autorisée par le responsable de la ZRR (ou son suppléant) qui peut interroger d'autres services compétents.

Il est rappelé que si les données partagées (dans le cadre d'une autorisation d'accès aux informations accordée à un tiers) se situent toujours au sein de la ZRR alors **il est obligatoire d'effectuer au préalable une demande d'autorisation d'accès comparable à celle d'un accès physique**.

En cas de doute, il convient de se référer à la politique de l'établissement et de contacter le fonctionnaire de sécurité défense.

Concernant la distinction entre accès et réutilisation des informations, cf. notion d'open data, p.5

1.3.3. Données protégées par le droit d'auteur et autres droits de propriété intellectuelle

1.3.3.1 Droit d'auteur

Lorsque les informations que je souhaite communiquer relèvent du droit d'auteur, je ne peux les communiquer (conférer un droit d'accès) que sous réserve de respecter les droits du(es) auteur(s)²². Par ailleurs, ces informations ne sont pas réutilisables par des tiers sans l'accord de(s) l'auteur(s).

Comment savoir s'il y a un droit d'auteur sur les informations que je souhaite diffuser ?

Les éléments ou données (textes, interviews, musiques ou sons, images, discours, représentations graphiques, etc.) :

- qui existent sous une forme concrète (leur existence est perceptible par les sens),
 - **ET** dont la forme est originale
- constituent une œuvre de l'esprit protégée par le droit d'auteur.

Seul l'auteur a le droit exclusif d'exploiter son œuvre, de façon directe ou par l'intermédiaire d'un tiers (*Exemple : par l'intermédiaire d'un éditeur*). L'auteur peut donc discrétionnairement autoriser ou interdire l'exploitation

Les idées exprimées dans un écrit peuvent être reprises librement. On dit qu'elles sont « de libre parcours ». En effet, le droit d'auteur protège la façon dont ces idées sont exprimées et non le contenu des idées.

La qualification d'œuvre de l'esprit pour des données photographiques recueillies de façon automatique (par exemple un appareillage) lors d'une expérimentation n'est pas tranchée par les tribunaux mais on peut penser qu'une photographie prise automatiquement par un appareillage n'est pas originale.

²¹ « Les ZRR sont, aux termes de l'article 413-7 du code pénal, constituées de locaux et de terrains clos dans lesquels l'accès et la circulation sont réglementés afin d'assurer la protection des installations, du matériel ou du secret des recherches, études ou fabrications ».

²² [Article L311-4 du Code des relations entre le public et l'administration](#)

de son œuvre et, en cas d'autorisation, il peut imposer des conditions d'exploitation.

Il existe une exception légale qui permet de reprendre certains éléments d'une œuvre sans autorisation de l'auteur. Il s'agit du droit de courte citation (article L122-5 Code de la Propriété Intellectuelle). J'ai donc le droit de citer une œuvre sous réserve :

- 1. que la citation soit courte et non substantielle par rapport à l'œuvre*
- 2. ET que soient cités le ou les auteur(s) et la source.*

Dans le contexte de l'Open Data des données de recherche, quatre cas principaux peuvent se présenter :

1) Les œuvres appartiennent aux scientifiques :

C'est le cas des œuvres réalisées par les chercheurs et enseignants-chercheurs.

Exemple : publications scientifiques.

Elles ne sont pas diffusables sans l'accord des scientifiques qui en sont les auteurs.

2) Les œuvres appartiennent à l'établissement public :

C'est le cas des œuvres réalisées par les agents autres que les chercheurs et enseignants-chercheurs.

C'est aussi le cas des œuvres dites « collectives » c'est-à-dire les œuvres commandées par l'établissement et réalisées selon les directives précises données par l'établissement.

Exemple : encyclopédie.

Elles sont diffusables par l'établissement sans accord préalable des auteurs dès lors que cette diffusion est strictement nécessaire à l'accomplissement d'une mission de service public²³.

Avertissement :

Les œuvres (textes, images, etc.) glanées sur internet sont soumises au droit d'auteur. Même si aucune indication n'existe sur le site, l'accord de l'auteur est indispensable préalablement à toute réutilisation.

3) Les œuvres appartiennent en tout ou partie à un tiers

a. Si ce tiers est un établissement public :

Si un citoyen me demande l'accès à une œuvre appartenant à un établissement public autre que mon établissement, je dois transférer sans attendre cette demande audit établissement.

Si je veux diffuser une œuvre appartenant à un autre établissement public, je dois lui demander de le faire.

Dans ces deux cas, lorsque l'œuvre a été réalisée dans le cadre d'une mission de service public, elle doit être diffusée. A contrario, si l'œuvre ne relève pas de la mission de service public de l'établissement, celui-ci peut fixer des conditions d'accès et de diffusion.

b. Si ce tiers est un établissement ou une personne privé(e) :

²³ [Article L131-3-1 Code de la Propriété Intellectuelle](#)

Si l'œuvre appartient à un tiers privé et a été réalisée en dehors d'une mission de service public, seul le tiers peut en autoriser l'accès et la réutilisation.

Son autorisation préalable est indispensable.

4) L'œuvre est un logiciel réalisé dans le cadre d'un établissement public

Quel qu'en soit l'auteur (chercheur, enseignant-chercheur, autre agent), si l'œuvre est un logiciel, c'est l'établissement public qui est propriétaire des droits d'auteur sur le logiciel²⁴.

En conséquence, c'est l'établissement qui décide de sa diffusion.

Dans tous les cas (1 à 4) : l'accord de l'auteur se fera impérativement sous forme de **licence écrite** indiquant les **supports de diffusion autorisés, la durée et les territoires couverts**. Ces conditions de forme sont rendues obligatoires par la loi²⁵.

La notion de service public est difficile à cerner car il n'existe pas de définition légale mais seulement des cas de jurisprudence. La qualification de service public dépend de la façon dont est exercée la mission et de l'objectif de cette mission (intérêt général). Exemple : France Télécom accompli certaines missions dans le cadre d'un service public (entretien du réseau de communication) et d'autres missions dans un cadre privé (vente de prestations, de matériels).

1.3.3.2 Autres droits de propriété intellectuelle

Lorsque les informations que je souhaite diffuser sont contenues dans des « documents » sur lesquels des tiers détiennent des droits de propriété intellectuelle²⁶, leur diffusion est soumise à l'accord préalable de ces tiers.

Ces droits de propriété intellectuelle regroupent les marques, les dessins et modèles, les brevets et le droit des producteurs de bases de données.

1. Marques, dessins et modèles :

Si les données que je veux diffuser contiennent des marques ou des dessins protégés, je ne peux pas les diffuser sans prendre certaines précautions et je dois prendre contact avec mon service juridique.

2. Brevets, éléments brevetés :

L'information sur le brevet est accessible au public au maximum 18 mois après le dépôt de la demande de brevet²⁷, donc rediffusable à compter de cette date.

Attention !! Reproduire l'information contenue dans le brevet ne signifie pas que j'ai le droit de reproduire le produit ou le procédé breveté. Seul est autorisée la vérification expérimentale que le brevet fonctionne (exemption de recherche). Toute autre reproduction de l'invention brevetée est interdite sans l'autorisation du titulaire du brevet et constitue une contrefaçon.

²⁴ [Article L113-9 code de la propriété intellectuelle](#)

²⁵ [Article L131-3 Code de la propriété intellectuelle](#)

²⁶ [Article L321-2 code des relations entre le public et l'administration](#)

²⁷ [Article L612-21 code de la propriété intellectuelle](#)

Pour en savoir plus sur les bases de données et l'identification du producteur :
Fiche 1 : Les bases de données en bref

3. Bases de données

Si les données sont structurées dans une base de données, celui qui souhaite accéder aux données a le droit de faire une **extraction non substantielle, soit qualitativement, soit quantitativement, s'il a licitement accès à la base**. Le producteur de la base de données ne peut pas s'opposer à ce droit²⁸.

En revanche, dès lors que l'extraction devient substantielle, l'accord du producteur est requis.

Le producteur est celui qui a fait les investissements financiers, matériels ou humains substantiels pour la création de la base et son enrichissement. Les tribunaux sont très sélectifs et ne prennent pas en compte les investissements réalisés pour créer les données. En revanche, comptent les moyens consacrés à la recherche d'éléments existants, à leur vérification et correction et à leur rassemblement dans la base.

Plusieurs cas peuvent se présenter :

- **le producteur de la base de données est mon établissement** : en principe, l'établissement doit communiquer la base de données au citoyen qui en fait la demande. Depuis la loi pour une République numérique, c'est un principe de diffusion plutôt que de communication sur demande qui doit s'appliquer. Pour connaître et analyser les conditions et limites de l'accès à la base et de réutilisation des données, je prends contact avec mon service juridique.
- **le producteur de la base de données est un autre établissement public** : il faut transmettre la demande d'accès du citoyen à cet établissement.
- **le producteur de la base de données est une personne physique ou morale de droit privé (hors mission de service public)** : il faut transmettre la demande d'accès du citoyen à cette personne qui pourra s'opposer à la réutilisation de sa base de données.

La loi pour une République numérique instaure, pour certaines administrations²⁹, des obligations importantes sur la communicabilité des données. Il convient donc de s'y référer avant toute diffusion d'une base de données.

²⁸ [Article L342-3 code de la propriété intellectuelle](#)

²⁹ notamment les collectivités territoriales de plus de 3 500 habitants.

1.3.4 Données relatives aux personnes, à la vie privée

Pour en savoir plus
Fiche 3 : Données personnelles

Si les informations que je souhaite diffuser sont relatives à des personnes physiques et permettent leur identification directe ou indirecte, alors il s'agit de données à caractère personnel³⁰.

Lors de la constitution d'un jeu de données qui va inclure des données personnelles, je ne peux collecter ni réaliser aucun traitement de ces données sans avoir, au préalable, respecté les « formalités CNIL »³¹ imposées par la loi 78-17 du 6 janvier 1978 modifiée.

Les formalités imposées par la loi diffèrent en fonction de la nature des données traitées et de la finalité poursuivie par le traitement (Cf. [Fiche n°3](#)).

Je dois être très vigilant vis-à-vis du respect des personnes et de leur vie privée³². Je dois me rapprocher du Correspondant Informatique et Libertés (CIL) ou, à défaut, des services juridiques de mon établissement lorsqu'une demande d'accès est formulée pour des documents ou jeux de données contenant des données à caractère personnel.

A condition d'avoir respecté les « formalités CNIL » et s'ils ne contiennent pas de données sensibles dont la collecte est prohibée, les documents comprenant des données à caractère personnel ne sont communicables et réutilisables que **si au moins l'une des trois conditions suivante est remplie** :

1. le consentement des personnes concernées a été recueilli après leur bonne information sur la finalité et les modalités de la communication ou de la réutilisation des données les concernant ;
2. les données sont anonymisées³³ (c'est-à-dire non identifiantes ou ne permettant pas, compte tenu de leur niveau d'agrégation, d'identifier à nouveau les personnes concernées) ;
3. la réutilisation est autorisée par un texte législatif ou réglementaire.

Lorsqu'une demande d'accès ou de réutilisation est formulée pour des données à caractère personnel communicables, la communication doit être encadrée par une licence adaptée aux risques de mauvaises réutilisations des données. La licence utilisée doit notamment garantir l'intégrité, la sécurité et la confidentialité des données transmises. Le

Attention !! La collecte de certaines données particulièrement sensibles est légalement prohibée sauf exceptions et sur autorisation de la CNIL ou adoption d'un décret en Conseil d'Etat. Exemples de données sensibles : origine raciale ou ethnique ; opinions politiques, syndicalistes, philosophiques ou religieuses ; données de santé, génétiques et biométriques.

Cas particulier des données des recherches en santé humaine : dans certains cas les données personnelles de santé sont communicables, uniquement à la personne concernée (patient) ou ses ayants droits, par l'intermédiaire d'un professionnel de santé qualifié (médecin) et sous certaines conditions. L'accès à ces données est contrôlé et restreint. Pour permettre la diffusion de ces données, il faut penser à leur anonymisation dès la collecte.

L'accès à la plupart des données médico-sociales est contrôlé par l'Institut des Données de Santé (IDS) Les demandes d'accès aux bases de données médico-administratives sont examinées par l'IDS au cas par cas et leur réutilisation est soumise à des conditions strictes. La CNIL doit être à même de vérifier l'efficacité des procédés d'anonymisation utilisés. Il est par ailleurs possible et recommandé de saisir la CNIL au préalable afin de faire valider un tel processus lorsqu'il concerne des traitements d'envergure.

³⁰ Article 2 loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés

³¹ Commission Nationale Informatique et Libertés (CNIL).

³² Article 9 du Code civil.

³³ Le processus d'anonymisation est en principe à la charge de l'administration qui détient les données demandées. D'après l'article 40 du décret n°2005-1755 du 30 décembre 2005 l'administration peut refuser d'anonymiser les données si cette opération entraîne pour elle des « efforts disproportionnés ».

demandeur doit en outre s'engager à ne pas tenter d'identifier à nouveau les personnes concernées par les données ayant été anonymisées.

Pour en savoir plus
[Fiche 4 : Données statistiques](#)

1.3.5 Données statistiques

Ces données proviennent des statistiques publiques qui regroupent l'ensemble des productions issues³⁴ :

- des enquêtes statistiques dont la liste est arrêtée chaque année par un arrêté du ministre chargé de l'économie ;
- de l'exploitation, à des fins d'information générale, de données collectées par des administrations, des organismes publics ou des organismes privés chargés d'une mission de service public.

Ces données peuvent contenir des données personnelles et des données sur les entreprises relevant du secret des affaires (parts de marché, etc.). Les données statistiques peuvent être diffusées à condition de respecter le « secret statistique » ou en obtenant une dérogation de la part du Comité du Secret Statistique.

L'exception relative aux prestations de services est peu documentée. Il n'existe pas d'avis de la CADA concernant cette exception appliquée aux établissements de recherche et d'enseignement supérieur.

1.3.6 Données liées à un contrat avec un tiers non soumis à une obligation de service public

Les règles exposées dans ce guide concernant la diffusion et la réutilisation des documents administratifs sont des règles d'ordre public qui prévalent sur les conditions que les parties peuvent fixer dans un contrat.

Pour rappel, et concernant les données acquises par l'établissement lors de relations avec des tiers privés, l'administration n'est pas tenue :

- de communiquer ou de diffuser les documents réalisés en exécution d'un contrat de prestation de services exécuté pour le compte d'une ou plusieurs personnes déterminées ; attention !! L'expression « prestation de service » est entendue au sens large : il s'agit de tous les contrats par lesquels l'établissement réalise des travaux (analyse, expérimentation, recherche, etc.) pour le compte d'un tiers. Cette exception est atténuée par le fait que, depuis la loi pour une République Numérique du 7 octobre 2016³⁵, il est prévu que la réutilisation de données issue d'une activité de recherche est libre si :

- ces données sont issues d'une activité de recherche financée au moins pour moitié par des fonds publics ; et

³⁴ Source : [INSEE](#).

³⁵ [Article L 533-4 II Code de la Recherche](#)

- ces données ne sont pas protégées par un droit spécifique ;
- et ces données ont été rendues publiques par le chercheur ou l'établissement.

- de communiquer ou de diffuser des données protégées par le secret des affaires (ou tout autre secret protégé par la loi) ;

- de conférer des droits de réutilisation sur des données sur lesquelles existent des droits de propriété intellectuelle de tiers.

Pour les données échangées ou recueillies lors d'un partenariat public-privé, il est conseillé d'établir un contrat détaillant l'usage que l'on peut faire des données.

Notamment, il est recommandé :

- de prévenir le partenaire, dans le contrat, que l'administration sera tenue de communiquer certaines données aux citoyens qui en feront la demande et que l'accès, voire la diffusion de certaines informations est une obligation légale pesant sur l'administration ;
- a minima de prévoir une exception à la confidentialité des informations échangées dans le cadre du partenariat ;
- de formuler un plan de gestion des données, qui abordera la problématique de la diffusion et de la réutilisation des données.

A titre d'exemple, le contrat UnicANR, utilisé dans le cadre de partenariats financés par l'ANR contient la clause suivante (qui préserve les facultés d'Open Data des partenaires publics) :

« Dans le cas où la communication d'INFORMATIONS CONFIDENTIELLES est imposée par l'application d'une disposition légale ou réglementaire ou dans le cadre d'une procédure judiciaire, administrative ou arbitrale, cette communication doit être limitée au strict nécessaire. La PARTIE RECIPIENDAIRE s'engage à informer immédiatement et préalablement à toute communication la PARTIE EMETTRICE afin de permettre à cette dernière de prendre les mesures appropriées à l'effet de préserver leur caractère confidentiel. »

La clause peut également indiquer que ne sont pas confidentielles les informations que la partie publique est légalement tenue de communiquer.

1.4 Conclusion

L'ensemble des développements de ce point 1 sont reportés sur la carte heuristique « Panorama de la communicabilité d'une donnée » que vous trouverez en Annexe.

2. Comment diffuser les données ?

2.1. Les grands principes :

2.1.1 Aspects techniques de l'Open Data

Les préconisations élaborées en 2007 par le groupe de travail [Open Government Data](#) définissent les bonnes pratiques relatives aux données mises en Open data. Ces données doivent être :

- Complètes et de qualité : données brutes accompagnées de leurs métadonnées.
- Accessibles et exploitables directement : format numérique non propriétaire, open source de préférence.

Ainsi, pour être considérées « ouvertes » au sens du mouvement « Open Data », les données publiques doivent être, dans la mesure du possible et sous réserve de certains choix laissés à l'appréciation des établissements publics :

- Complètes : toutes les données sont mises à disposition, à l'exception des données qui sont sujettes à des limitations. *Exemple : données personnelles.*
- Primaires : les données sont ouvertes telles que collectées à la source, avec le moins de modifications et la plus grande granularité possibles.
- Opportunes : les données sont mises à disposition aussi rapidement que nécessaire pour préserver leur valeur.
- Accessibles : les données sont accessibles au plus grand éventail d'utilisateurs possible et pour des usages aussi divers que possibles.
- Exploitables : les données sont exploitables par ordinateur ou lisibles par les machines. Les données sont structurées pour permettre le traitement automatisé.
- Non-Discriminatoires : les données sont non-discriminatoires, c'est à dire accessibles à tous, sans aucune obligation préalable ni inscription et en principe sans coût.
- Non-Propriétaires : les données sont accessibles dans un format sur lequel aucune entité ne dispose d'un contrôle exclusif. *Exemple : format de fichier texte non propriétaire permettant une réutilisation facile et ne nécessitant pas l'achat d'une licence.*
- Libres de droits : les données sont libres de droit. Elles ne sont pas soumises au droit d'auteur, au droit des marques ou au secret commercial. Des règles raisonnables de confidentialité, de sécurité et de priorité d'accès peuvent être admises.
- Permanentes : les données doivent être rendues accessibles en ligne avec un système d'archivage, de suivi des modifications et de mention de la dernière mise à jour.

La réglementation fait écho à certaines de ces recommandations en obligeant à la diffusion de documents achevés et en recommandant que la diffusion électronique soit effectuée dans un standard ouvert et librement réutilisable.

Bien entendu, l'application des recommandations purement techniques est laissé à l'appréciation de l'établissement.

2.1.2 Aspects juridiques : choix d'une licence de diffusion

Quelle licence ?

Si le contrôle de la nature des données a démontré la possibilité de leur diffusion en Open Data, il reste un dernier point à préparer : le choix d'une licence.

L'établissement a le choix entre plusieurs licences et peut en préconiser certaines. Cependant, la Loi pour une République Numérique du 7 octobre 2016 indique que la licence doit être choisie sur une liste de licences fixée par décret ou, à défaut, selon une licence homologuée par l'Etat. Les décrets fixant ces licences et la procédure d'homologation ne sont pas encore publiés.

Tout « réutilisateur » peut faire l'usage qu'il souhaite des données, sous réserve de³⁶ :

- **respecter l'intégrité des données** (absence d'altération, absence de dénaturation) ;
- **faire mention de la source des données** ;
- **veiller à ce que l'indication de la date de dernière mise à jour soit bien présente.**

L'imposition d'une licence à l'utilisateur permet d'exclure la responsabilité directe du fournisseur des données lors de la réutilisation (clause de limitation ou d'exclusion de responsabilité).

Si les données ont été obtenues au cours d'un partenariat, il faut vérifier le contrat de partenariat pour savoir si l'établissement peut prendre seul ou non le choix de la licence. Si ce choix doit être réalisé avec le partenaire, il faut le contacter.

Il existe de nombreuses licences libres et gratuites pour répondre aux besoins d'une ouverture de données de recherche dans le cadre de l'Open Data. Les plus courantes sont les licences Creative Commons 4.0 (CC), l'Open Data base License (ODbL) de l'Open Knowledge Foundation et la licence ouverte Etalab (initiative gouvernementale).

Attention !!
Certaines versions de licences Creative Commons sont plus permissives que d'autres.

*Pour en savoir plus sur les licences Creative Commons (CC) 4.0
Fiche 5 : Tableau comparatif des licences CC 4.0
Fiche 6 : Tableau comparatif des licences gratuites ODBL et Etalab*

³⁶ [Article L322-1 Code des relations entre le public et l'administration](#)

Licence gratuite ou payante ?

Depuis la loi dite Valter du 28 décembre 2015³⁷, c'est le principe de gratuité de la réutilisation qui est consacré.

L'accès payant aux données n'est désormais possible que que les établissements qui sont tenus de couvrir par des recettes propres une part substantielle des coûts liés à l'accomplissement de leurs missions de service public³⁸.

Exemple : Météo France, INSEE

Pour savoir si votre établissement est concerné, vous pouvez contacter le service juridique.

2.2. Les différentes modalités de diffusion :

L'ouverture des données peut être volontaire ou « forcée » par la demande d'un tiers, c'est pourquoi il est proposé d'aborder les modalités d'ouverture suivantes : prise en compte de pratiques par discipline ([2.2.1](#)) et existence d'une politique d'établissement ([2.2.2](#)). Dans tous les cas, l'établissement doit mettre en place un répertoire recensant les données publiques à disposition du citoyen ([2.2.3](#)). Enfin, nous parlerons des réflexes à avoir lors d'une demande d'accès formulée par un tiers ([2.2.4](#)).

Attention !!

Je ne dois pas choisir d'entrepôt de données par moi-même sans me conformer à la politique de mon établissement ou sans m'être rapproché de ma hiérarchie. En effet, certains entrepôts ou hébergeurs peuvent avoir des conditions d'utilisation de nature à confisquer le droit d'exploitation ultérieure des données ou imposer des restrictions de réutilisation des données qui seraient incompatibles avec le droit français.

2.2.1. Les pratiques par discipline

Certaines disciplines scientifiques proposent, des entrepôts reconnus permettant l'hébergement des jeux de données (ex : NCBI, EBI pour la bioinformatique) pour leur mise en accès libre notamment à la communauté scientifique. Ces entrepôts disciplinaires peuvent tout à fait constituer une façon de rendre accessible les données.

Toutefois, l'obligation d'open data faite aux établissements publics par la réglementation prescrit d'ouvrir les données à **tout public** (citoyen, entreprise, organisme d'origine française ou étrangère). Il faut donc vérifier que l'entrepôt de données ne restreint pas l'accès aux seuls scientifiques.

Si c'est le cas, il convient d'envisager un dépôt complémentaire (cf. ci-dessous) dans un entrepôt totalement ouvert.

³⁷ [Loi 2015-1779](#)

³⁸ [Décret 2016-1036 du 28/07/2016](#)

2.2.2. La politique d'Etablissement

Afin de répondre aux enjeux de l' « Open » et du « Big » data, un établissement peut se doter d'une offre de services en direction des équipes de recherche pour les aider à rentrer dans le mouvement de l'open data. A l'occasion de cette démarche de création de services en faveur du partage des données, les établissements peuvent mettre en place :

- un annuaire des sources de données ;
- un entrepôt de données : également appelé « data repository », il s'agit d'un réservoir constitué majoritairement de données de recherche, brutes ou élaborées, qui sont organisées de façon à pouvoir être retrouvées.

Les EPIC n'ont pas l'obligation légale d'autoriser la réutilisation des données qui ne rentrent pas dans le cadre de leur mission de service public.

Ce type d'entrepôt devrait être en mesure d'héberger des données provenant de divers projets ou de fédérer des systèmes d'informations, projets ou plateformes³⁹.

A défaut d'entrepôt d'établissement disponible, je peux ouvrir les données en les mettant à disposition dans un entrepôt de référence de données conformément à la politique de mon établissement.

2.2.3. Le répertoire des données publiques

La réglementation oblige chaque établissement à disposer d'un répertoire des informations publiques (= documents publics réutilisables). Tout dépôt de données en open data doit donc être mentionné dans ce répertoire.

Dans tous les cas il est fortement recommandé, lors de l'ouverture des données, de mettre en place une licence.

Je dois me renseigner auprès de mon établissement pour connaître le répertoire des informations publiques (Exemple : à l'INRA, il s'agit de ProdINRA).

2.2.4. La demande d'accès formulée par un tiers

Si je suis contacté par un tiers (citoyen, association, autre) qui demande accès à certaines données que j'ai créées ou collectées, je dois réagir sans attendre.

La loi oblige à formuler une réponse au tiers dans un délai d'un mois⁴⁰.

³⁹ Open Science, Groupe de travail INRA données expérimentales

⁴⁰ [Article R311-13 du code des relations entre le public et l'administration](#)

Faute de réponse dans ce délai, l'accès aux informations est refusé.

Le demandeur peut alors exercer un recours devant la CADA (Commission d'Accès aux Documents Administratifs). La CADA rend un avis (qui sera publié sur son site) qui n'oblige pas l'établissement.

Si l'établissement décide de ne pas se conformer à l'avis, le demandeur peut exercer un recours devant la juridiction administrative.

En cas de demande d'un tiers, je dois donc contacter immédiatement le service juridique de mon établissement qui m'aidera à formaliser la réponse.

Sous réserve de la communicabilité des données (cf 1), je peux me contenter de fournir au tiers demandeur les informations en l'état. Je n'ai pas l'obligation de les retraiter ou encore de modifier le format dans lequel les données sont disponibles (mais la loi précise que les données sont mises à disposition dans un format ouvert et aisément réutilisable⁴¹). Toutefois, dans la mesure du possible les modalités de mise à disposition et les formats doivent favoriser l'exploitation des informations par les utilisateurs.

Si des données communicables sont mêlées à des données non communicables, telles que des données personnelles, et s'il est possible de disjoindre les données communicables et non communicables ou d'occulter les données non communicables, alors il faut répondre positivement à la demande d'accès aux données⁴³, en précisant ses limites.

Dans ce cas et dans l'hypothèse où les données ne sont pas communicables, le service juridique de mon établissement m'aidera à faire une réponse qui se doit de respecter un certain formalisme et d'être circonstanciée.

Novembre 2016

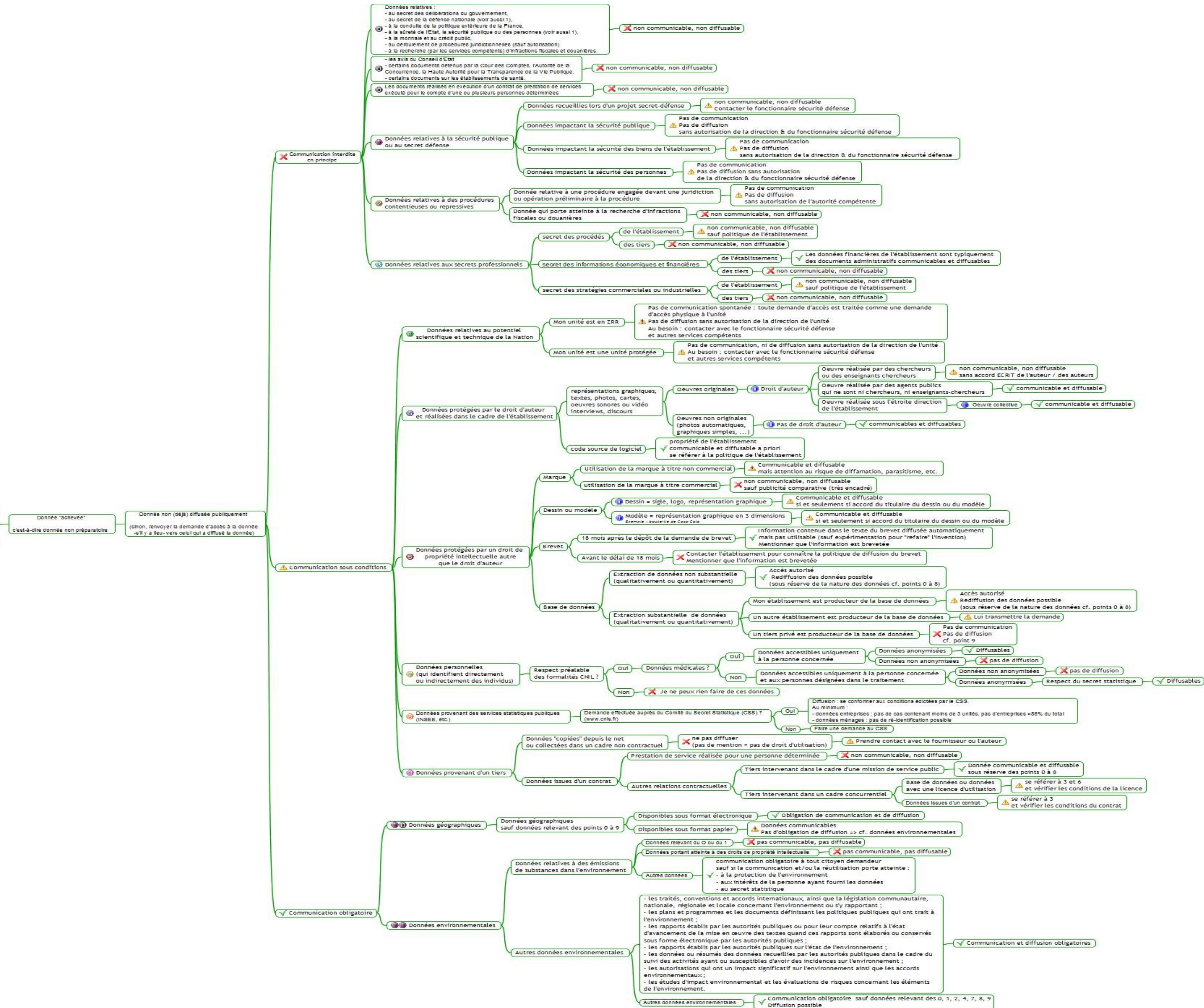
⁴¹ Article 10 loi n°78-753 dite « CADA » modifié par la loi n°2015-1779 du 28.12.2015

⁴² Pour en savoir plus sur le régime juridique applicable en cas de détention ou codétention d'une information dans l'exercice d'une mission de service public administratif (SPA) et/ou d'une mission de service public industriel et commercial (SPIC) ; consultez le [« Cahier pratique : Ressources de l'immatériel »](#) de l'APIE, page 8.

⁴³ Article L 311-7 Code des relations entre le public et l'administration

Logigramme de communicabilité d’une donnée,

Cf. Carte heuristique



Fiche 1 : Les bases de données en bref

La protection des bases de données relève de deux droits différents : le droit d'auteur et le droit « sui generis » du producteur de base de données. Ce dernier a été instauré par la directive 96/9/CE du Parlement européen et du Conseil du 11 mars 1996, transposée par la loi n° 98-536 du 1er juillet 1998. Le Code de la propriété intellectuelle définit depuis lors la notion de base de données comme un « recueil d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessibles par des moyens électroniques ou par tout autre moyen »⁴⁴.

Le droit d'auteur portant sur une base de données naît dès lors que sa structure est originale, c'est-à-dire que la disposition des éléments qu'elle inclut relève d'un choix qui « reflète l'empreinte de la personnalité de l'auteur ». Ainsi, un annuaire listant des personnes par ordre alphabétique ne caractérise pas une structure originale.

Devant les faiblesses du droit d'auteur, qui ne permet pas de reconnaître de protection aux bases de données ayant fait l'objet d'investissements lourds mais dont la structure n'est pas originale, le législateur a fondé le droit « sui generis ».

L'objectif de cette législation est d'encourager le développement des bases de données en instaurant un droit spécifique ; le droit sui generis, destiné à protéger les investissements financiers, matériels et humains consentis par les producteurs de bases de données. Ce droit protège avant tout le contenu informationnel de la base, indépendamment de la protection par le droit d'auteur. L'investissement doit néanmoins être substantiel, que ce soit quantitativement ou qualitativement.

La directive 96/9/CE cherche en effet à protéger l'investissement substantiel nécessaire à l'obtention, la vérification ou la présentation des données de la base.

- Obtention : « moyens consacrés à la recherche d'éléments existants et à leur rassemblement dans ladite base. Elle ne comprend pas les moyens mis en œuvre pour la création des éléments constitutifs du contenu d'une base de données »⁴⁵.
- Vérification : « moyens consacrés, en vue d'assurer la fiabilité de l'information contenue dans ladite base, au contrôle de l'exactitude des éléments recherchés, lors de la constitution de cette base ainsi que pendant la période de fonctionnement de celle-ci. Des moyens consacrés à des opérations de vérification au cours de la phase de création d'éléments par la suite

Note importante

Si l'article L111-1 du Code de la propriété intellectuelle reconnaît la qualité d'auteur au fonctionnaire, le droit d'exploitation d'une œuvre créée par un agent de l'Etat est le plus souvent « cédé de plein droit à l'Etat » (article L131-3-1 du même code). Ce n'est en revanche pas le cas, d'après l'alinéa 4 de l'article L111-1, pour les « agents auteurs d'œuvre dont la divulgation n'est soumise, en vertu de leur statut ou des règles qui régissent leurs fonctions, à aucun contrôle préalable de l'autorité hiérarchique », ce qui vise notamment les maîtres de conférences, les professeurs d'universités, les chargés de recherches ou encore les directeurs de recherches. Ces agents ont pu être qualifiés, à l'occasion de débats parlementaires, d'« agents qui disposent dans leurs fonctions d'une grande autonomie intellectuelle, voire une indépendance de jugement, même si celle-ci s'inscrit dans une hiérarchie ».

Par principe, le droit d'exploitation de l'œuvre créée par un agent est dévolu à l'Etat. Par exception pour la catégorie d'agents indiquée, le droit d'exploitation est conservé par l'agent auteur à qui il faut alors faire signer une cession de droits.

⁴⁴ Article L112-3 du Code de la Propriété Intellectuelle (CPI).

⁴⁵ CJCE, 9 nov. 2004, The British Horseracing Board Ltd e/a.

rassemblés dans une base de données ne relèvent pas de cette notion »⁴⁶.

- Présentation : « moyens visant à conférer à ladite base sa fonction de traitement de l'information, à savoir ceux consacrés à la disposition systématique ou méthodique des éléments contenus dans cette base ainsi qu'à l'organisation de leur accessibilité individuelle »⁴⁷.

Si les données dont je dispose sont structurées dans une base de données, je dois identifier le producteur de cette base pour pouvoir procéder à l'ouverture des données. En effet, le producteur de la base de données est le seul à pouvoir autoriser l'ouverture. Pour l'identifier, je dois me poser une série de questions :

❖ Qui a pris l'initiative du projet ?

- Contrat de consortium, de recherche ou de collaboration : prendre connaissance des termes du contrat relatif au projet, et prêter une attention particulière aux clauses de propriété intellectuelle.
- Collaboration informelle : que ce soit entre équipes ou intra-équipe de recherche, il faudra recenser l'ensemble des employeurs et acteurs pour déterminer les participations (matérielles, intellectuelles ou financières) de chacun.

❖ Qui a réalisé l'architecture de la base de données ?

- Contrat avec une société privée (ESN ou SS2i) : contrôler la clause de propriété intellectuelle pour savoir si le prestataire reste propriétaire de la structure .
- Typologie de l'architecture : est-elle d'une conception classique pour une base de données, ou est-elle plutôt originale par rapport à ce qui existe ? Si oui, l'auteur de la structure a-t-il cédé les droits nécessaires à son utilisation ?
- Base de données composite : si la base de données est composite et a été constituée par l'agrégation, même partielle, d'autres bases de données, il faudra respecter les licences d'utilisation des bases de données sur laquelle elle s'appuie et veiller à détenir les cessions de droits nécessaires à l'utilisation des structures préexistantes.

❖ Et donc, qui est le producteur de la base de données ?

⁴⁶ CJCE , 9 nov. 2004 The British Horseracing Board Ltd e/a précité.

⁴⁷ CJCE, 9 nov. 2004, Fixtures Marketing Ltd c/ Organismos Pronostikon.

- Une fois l'ensemble des acteurs déterminés⁴⁸, la solution la plus simple pour savoir qui aura des droits et à qui demander l'autorisation d'ouvrir la base de données serait d'attribuer, dès le début de la collaboration soit :
 - une part égale à chaque acteur : tous seront alors dans un régime proche de l'indivision, c'est-à-dire avec nécessité de concertation et unanimité pour prendre la plupart des décisions sauf à avoir organisé l'exploitation de la base selon des règles différentes à la manière d'un règlement de copropriété.
 - une part différente pour chacun : sur la base des investissements de chacun.

Mais des négociations peuvent intervenir a posteriori entre les acteurs et les débats auront lieu sur le terrain de la preuve des investissements de chacun.

Dans tous les cas, il est très important de conserver des preuves de la nature et de l'importance de l'investissement réalisé en concevant un dossier bien documenté de tout élément ayant servi à l'investissement de la base de données :

- Investissement financier : coût de la base de données (factures, contrats, etc.).
- Investissement humain : le nombre de personnes chargées de la création et/ou de la mise à jour de la base de données (contrats de travail).
- Investissement matériel : équipements nécessaires au développement de la base de données (factures).

Attention : il ne faut prendre en compte que les investissements portant sur l'obtention, la vérification ou la présentation de la base, et non pas ceux consentis pour la création des données. *Exemple : les investissements consentis pour mieux organiser la base, pour la rendre plus intelligible, sont pris en compte.*

Le producteur de la base de données a le droit d'interdire :

1. « L'extraction, par transfert permanent ou temporaire de la totalité ou d'une partie qualitativement ou quantitativement

⁴⁸ Rappel : uniquement les acteurs ayant eu un rôle d'investissement pour l'obtention, la collecte, la vérification ou la présentation de la base de données.

Attention : il existe un arrêt de Cour d'Appel qui a jugé que le producteur de la base de données ne peut exercer ces droits d'interdiction qu'à condition de les avoir mentionnés lors de la mise à disposition de la base¹.

Dès lors, il faudrait mentionner de façon claire et évidente à côté de chaque base de données publiée, que le droit du producteur de bases de données est expressément réservé.

Exemple : licence interdisant l'extraction et la réutilisation sans autorisation expresse préalable.

Cet arrêt a fait l'objet d'un pourvoi en cassation¹ qui a confirmé cette interprétation, mais qui reste critiqué.

En effet, on peut douter de la portée d'une telle décision : celle-ci traitait du volet pénal de la protection du droit du producteur, qui implique d'opérer une interprétation stricte de la loi. Par ailleurs, imposer la mention expresse d'une interdiction de la part du producteur pour mettre en œuvre son droit priverait la protection légale d'une grande partie de son fondement

substantielle du contenu d'une base de données sur un autre support, par tout moyen et sous toute forme que ce soit ;

2. La réutilisation, par la mise à la disposition du public de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu de la base, quelle qu'en soit la forme »⁴⁹.

Depuis la loi pour une République Numérique du 7 octobre 2016⁵⁰, il est prévu que la réutilisation de données issue d'une activité de recherche est libre si :

- ces données sont issues d'une activité de recherche financée au moins pour moitié par des fonds publics ; et
- ces données ne sont pas protégées par un droit spécifique ;
- et ces données ont été rendues publiques par le chercheur ou l'établissement.

Cette ouverture de la base de données par le producteur de la base de données devra également s'accompagner de la cession des droits du ou des auteur(s) de la structure de la base de données. En effet, l'utilisation et la diffusion de l'architecture, et donc de la base de données, sont soumises à l'autorisation préalable de l'auteur qui conservera dans tous les cas des droits moraux de paternité et d'intégrité sur son œuvre. Le concepteur de l'architecture de la base peut être à la fois coproducteur de l'ensemble de la base et auteur unique de son architecture.

⁵⁰ Article L 533-4 II Code de la Recherche

Fiche 2 : Les données personnelles

1-Définitions

D'après l'article 2 de la loi n°78-17 « Informatique et libertés » du 6 janvier 1978, « constitue une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne ». Par ailleurs, « la personne concernée par un traitement de données à caractère personnel est celle à laquelle se rapportent les données qui font l'objet du traitement ».

Les données sont donc considérées « à caractère personnel » dès lors qu'elles concernent des personnes physiques :

- Identifiées directement : lorsque par exemple son nom apparaît dans un fichier.
- Identifiables indirectement : lorsqu'un fichier comporte des informations telles que l'adresse I.P., le n° d'immatriculation, le n° de téléphone, un numéro d'identification lié à un fichier où se trouvent les données à caractère personnel, une photographie, des éléments biométriques, etc.

Pour déterminer si une personne est identifiable via les données traitées, il faut donc analyser les risques en fonction du contexte et des moyens à disposition des utilisateurs leur permettant d'identifier cette personne. Par exemple, un croisement de données peut permettre une identification indirecte de la personne concernée : « Certaines données peuvent donc constituer des données à caractère personnel si elles permettent d'identifier indirectement ou par recoupement d'informations une personne précise. Il peut en effet s'agir d'informations qui ne sont pas associées au nom d'une personne mais qui permettent aisément de l'identifier et de connaître ses habitudes ou ses goûts »⁵¹. Exemple : une date de naissance associée à une commune de résidence.

Les principes de protection des données à caractère personnel s'appliquent à toute opération ou tout ensemble d'opérations (appelés « traitement ») portant sur de telles données, quelque soit le procédé utilisé, et notamment à la collecte, l'enregistrement, l'organisation, la

ATTENTION

Le règlement général européen sur la protection des données à caractère personnel sera applicable le 15 avril 2018. Cette fiche sera revue en conséquence. D'ici là, les règles applicables restent celles décrites ci-dessous.

Exemple : une date de naissance associée à une commune de résidence

⁵¹ [Commission Nationale Informatique et Libertés \(CNIL\)](#)

conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication quelque soit la forme de mise à disposition, le rapprochement ou interconnexion, ainsi que le verrouillage, l'effacement ou la destruction.

Un traitement ne peut être mis en place qu'après avoir accompli les formalités imposées par la loi « Informatique et Libertés » et ce, sous réserve du respect de la confidentialité et de la sécurité des données collectées. Les données à caractère personnel ne sont pas, en tant que telles, éligibles à la mise à disposition en open data. Il est nécessaire d'avoir, au préalable, obtenu le consentement éclairé (sur la finalité et les modalités de la communication des données) de la personne concernée ou bien d'avoir anonymisé⁵² les données.

2-Principes

La loi « Informatique et Libertés » définit les principes à respecter lors de la collecte, du traitement et de la conservation des données personnelles :

- Principe de finalité : les données à caractère personnel ne peuvent être recueillies et traitées que pour un usage légitime et déterminé correspondant aux missions de l'établissement.
- Principe de proportionnalité : seules doivent être enregistrées les informations pertinentes et nécessaires vis-à-vis de la finalité déclarée.
- Principe de durée limitée de conservation des données : les informations ne peuvent être conservées de façon indéfinie dans les fichiers informatiques. Une durée de conservation doit être établie en fonction de la finalité de chaque fichier. Par la suite les données doivent être supprimées ou archivées sur un support distinct (car elles ne doivent plus être utilisées). Il y a toutefois quelques exceptions notamment pour les données conservées « en vue d'être traitées à des fins historiques, statistiques ou scientifiques »⁵³.
- Principe de sécurité et de confidentialité : le responsable du traitement⁵⁴ est astreint à une obligation de sécurité. Il doit prendre les mesures nécessaires pour garantir la confidentialité, l'intégrité et la sécurité des données.
- Principe de transparence : le responsable du traitement doit informer les personnes des traitements auxquels leurs données sont soumises tout en leur accordant un droit d'accès, de

⁵² Cf. la fiche n°16 [du guide CNIL sur la Sécurité des Données Personnelles](#)

⁵³ [Article 36 alinéa 1](#) de la loi 78-17 du 6 janvier 1978 modifiée.

⁵⁴ Le responsable du traitement de données à caractère personnel est la personne, l'autorité publique, le service ou l'organisme qui détermine ses finalités (à quoi il va servir) et ses moyens (selon quelles modalités).

modification, de rectification voire d'opposition au traitement. Le responsable du traitement doit avertir ces personnes dès la collecte des données et en cas de transmission de ces données à des tiers.

Certaines données dites « sensibles » bénéficient d'un régime spécifique. Les données sensibles sont celles qui font apparaître, directement ou indirectement :

- **les origines raciales ou ethniques ;**
- **les opinions politiques, philosophiques ou religieuses ;**
- **l'appartenance syndicale des personnes ;**
- **des informations relatives à la santé ou à la vie sexuelle.**

Par principe, la collecte et le traitement de ces données sont interdits. Cependant, dans la mesure où la finalité du traitement l'exige, ne sont pas soumis à cette interdiction :

- les traitements pour lesquels la personne concernée a donné son consentement exprès ;
- les traitements justifiés par un intérêt public après autorisation de la CNIL ou adoption d'un décret en Conseil d'Etat.

Certaines autres données à risque doivent être traitées avec les mêmes égards : données génétiques, données relatives aux infractions pénales, aux condamnations, données comportant des appréciations sur les difficultés sociales des personnes, données biométriques, données comprenant le numéro NIR⁵⁵, etc.

3-Formalités préalables au traitement initial

Tout fichier ou traitement informatisé comportant des données personnelles doit être déclaré au Correspondant informatique et libertés (CIL) qui, selon le type de données ou de finalité du traitement, l'inscrit au registre des traitements de l'établissement ou instruit avec le responsable de traitement la demande d'autorisation auprès de la CNIL.

Pour les traitements de données à caractère personnel courantes (cas des données ne relevant pas des données sensibles ou d'autres données à risque comme les données génétiques, les infractions pénales, les données comportant des appréciations sur les difficultés sociales des personnes, les données biométriques, les données comprenant le numéro NIR) le CIL est habilité à donner l'autorisation de la mise en œuvre du traitement et à en superviser le déroulement dans le respect de la loi.

⁵⁵ Numéro d'inscription des personnes (NIR) : également appelé numéro INSEE ou numéro de sécurité sociale.

*Pour en savoir plus sur les transferts hors UE
Guide de la CNIL sur « les transferts de données à caractère personnel hors Union Européenne »*

Pour les transferts de données vers les USA, il existe un accord spécifique entre l'Union Européenne et les USA intitulé le « privacy shield »

*Pour en savoir plus
Rapprochez-vous du C.I.L.
de votre établissement*

Dans l'hypothèse d'un transfert de données hors de l'Union Européenne (UE) et hors pays figurant sur la liste de protection adéquate, il est très important d'avertir au préalable le CIL ou, à défaut, d'effectuer une déclaration préalable auprès de la CNIL.

Pour certaines catégories de traitements, soit en raison des typologies de données enregistrées (ex : données sensibles), soit en raison de leurs finalités spécifiques ou des risques qu'ils comportent, la loi a prévu des formalités particulières d'autorisation préalable délivrée par la CNIL⁵⁶. Cette autorisation peut intervenir, dans certains cas, après consultation et avis d'un Comité d'experts. *Exemple : traitements de données à caractère personnel à des fins de recherche en santé ou d'évaluation des pratiques de soins.*

Certains traitements mis en œuvre par des organismes publics doivent donc recueillir l'autorisation préalable de la CNIL. Cette procédure concerne les traitements mis en œuvre par des organismes publics ou des organismes privés gérant un service public et qui concernent :

- la recherche en santé (exception faite des recherches couvertes par une Méthodologie de Référence CNIL [MR001 / MR002] et qui en respectent scrupuleusement les dispositions par la voie d'un engagement de conformité auprès de la CNIL) ;
- l'utilisation du NIR (n° de sécurité sociale) ou la consultation du RNIPP (lorsque les organismes ne sont pas déjà habilités) ;
- l'utilisation de données biométriques (empreintes digitales, contour de la main, iris de l'œil, etc.) ;
- le recensement de la population ;
- les téléservices de l'administration électronique ;
- la sûreté, la défense ou la sécurité publique ;
- la prévention, la recherche, la constatation ou la poursuite d'infractions pénales ou l'exécution des condamnations pénales ou des mesures de sûreté ;

Quand les principes et formalités ont été respectés, la question de la mise en open data des données collectées peut être instruite (cf 2.1.5 du présent guide).

⁵⁶ [Articles 26 et 27 de la loi du 6 janvier 1978 modifiée.](#)

Fiche 3 : Les données statistiques

Le secret statistique est une forme particulière du secret professionnel qui s'applique aux statisticiens, chargés de recueillir et d'exploiter des statistiques publiques. Son principe général est d'apporter aux personnes qui fournissent des informations utilisées pour l'établissement de statistiques publiques l'assurance que ces informations ne seront pas utilisées d'une façon susceptible de leur porter atteinte⁵⁷.

En France, le secret statistique est garanti par la loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques. Au niveau européen, la confidentialité des informations statistiques est affirmée dans plusieurs documents⁵⁸.

Le secret statistique permet d'assurer :

- aux personnes physiques que la confidentialité sur leur vie personnelle et familiale sera garantie ;
- aux entreprises que le secret commercial sera respecté : les informations transmises ne seront pas mises à la disposition de leurs concurrents.

Les données des enquêtes statistiques sont détenues par l'INSEE et les services ministériels chargés des questions statistiques.

Toutefois, les données des enquêtes publiques françaises peuvent être accessibles à des fins scientifiques, après instruction d'une demande formulée auprès du Comité du Secret Statistique. Pour veiller à l'anonymat des données agrégées qui sont publiées après retraitement, les recommandations suivantes sont à respecter :

- Aucune donnée publiée ne doit concerner moins de 3 entités à la fois ;
- Aucune donnée publiée ne doit concerner une seule entreprise pour plus de 85% du total ;
- Aucune donnée publiée ne doit permettre l'identification directe ou indirecte des personnes.

L'ensemble des recommandations et interdictions (par exemple : prohibition de diffusion de données individuelles fiscales) est disponible dans le guide du secret statistique édité par l'INSEE.

Pour ce qui concerne les données collectées au niveau communautaire par les autorités européennes, le règlement n° 322/97 organise

⁵⁷ Source : [INSEE - Guide du secret statistique](#).

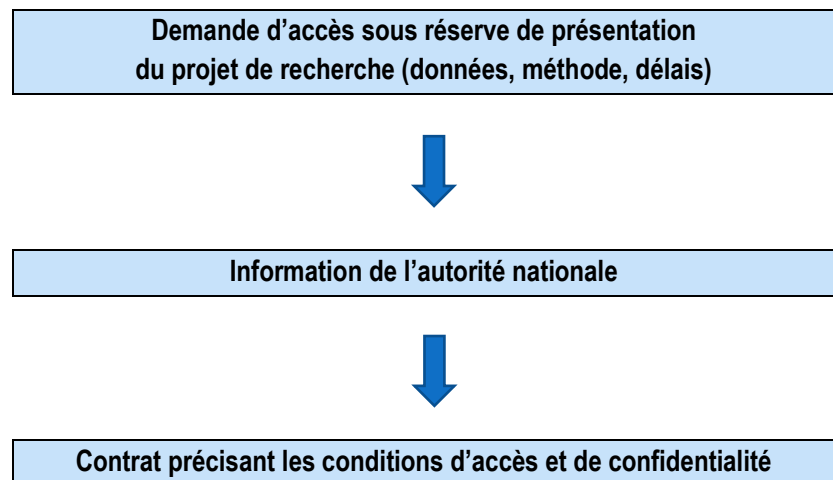
⁵⁸ [Article 338 du Traité UE](#) ; [Règlement n° 322/97 du Conseil du 17 février 1997](#) ; [Règlement d'application n° 831/2002](#)

l'établissement de statistiques européennes avec deux sources possibles :

- les données collectées par les états membres sont transmises aux autorités européennes ;
- et/ou les autorités européennes organisent des enquêtes.

Il existe un principe de diffusion des résultats statistiques obtenus.

En parallèle de la réglementation française, existent des modalités d'accès aux données collectées par l'Union Européenne à des fins scientifiques. C'est le règlement n°831/2002 qui s'applique. Celui-ci prévoit une procédure similaire à celle qui existe en France :



Fiche 4 : Convention d'Aarhus sur l'information en matière d'environnement

La convention d'Aarhus (Danemark) est une convention internationale signée le 25 juin 1998 et entrée en vigueur en octobre 2001. Cette convention concerne l'accès à l'information, la participation du public au processus décisionnel et l'accès à la justice en matière d'environnement.

La convention pose trois principes :

- 1) le développement de l'accès au public en matière d'information environnementale détenue par les autorités,
- 2) la participation du public aux décisions relatives à l'environnement,
- 3) l'accès à la justice du public pour faciliter les recours en cas de non respect ou de contestation du respect des deux premiers principes.

Cette convention a été transposée en droit européen puis en droit français⁵⁹.

L'effet majeur de cette convention est que certaines informations doivent faire l'objet d'une diffusion publique pour que les citoyens puissent exercer leurs droits de participation et de recours en matière d'environnement.

Pour pouvoir appliquer ces droits, il s'est avéré nécessaire de définir la notion d'information en matière d'environnement dans le Code de l'Environnement ([article L124-2](#)). Il s'agit de toute information disponible, quel qu'en soit le support, qui a pour objet :

- l'état des éléments de l'environnement, notamment l'air, l'atmosphère, l'eau, le sol, les terres, les paysages, les sites naturels, les zones côtières ou marines et la diversité biologique, ainsi que les interactions entre ces éléments ;
- les décisions, les activités et les facteurs, notamment les substances, l'énergie, le bruit, les rayonnements, les déchets, les émissions, les déversements et autres rejets ;
- l'état de la santé humaine, la sécurité et les conditions de vie des personnes, les constructions et le patrimoine culturel, dans la mesure où ils sont ou peuvent être altérés par des éléments de l'environnement, des décisions, des activités ou des facteurs mentionnés ci-dessus ;

⁵⁹ [Décret 2002-1187 du 12 septembre 2002](#)

- les rapports établis par les autorités publiques ou pour leur compte sur l'application des dispositions législatives et réglementaires relatives à l'environnement.

En matière d'accès à l'information environnementale par le public, il existe deux modalités :

- soit l'autorité publique diffuse l'information ;
- soit l'autorité publique répond aux demandes des citoyens.








Sur le premier point, le droit français indique que les informations suivantes doivent faire l'objet d'une diffusion publique :

- les traités, conventions et accords internationaux, ainsi que la législation communautaire, nationale, régionale et locale concernant l'environnement ou s'y rapportant ;
- les plans et programmes et les documents définissant les politiques publiques qui ont trait à l'environnement ;
- les rapports établis par les autorités publiques ou pour leur compte relatifs à l'état d'avancement de la mise en œuvre des textes quand ces rapports sont élaborés ou conservés sous forme électronique par les autorités publiques ;
- les rapports établis par les autorités publiques sur l'état de l'environnement ;
- les données ou résumés des données recueillies par les autorités publiques dans le cadre du suivi des activités ayant ou susceptibles d'avoir des incidences sur l'environnement ;
- les autorisations qui ont un impact significatif sur l'environnement ainsi que les accords environnementaux ;
- les études d'impact environnemental et les évaluations de risques concernant les éléments de l'environnement.

Concernant les informations environnementales qui seraient contenues dans d'autres documents, la convention d'Aarhus encourage les états à les diffuser au public. Toutefois, la transposition de cette convention dans le droit français⁶⁰ indique que ces informations sont communicables suivant le même régime que les documents administratifs (les exceptions à l'accès et à la réutilisation des documents sont quasiment les mêmes que pour les autres catégories de données publiques), sauf pour les données relatives à l'émission de substances dans l'environnement pour lesquelles les restrictions d'accès sont très réduites (autrement dit ces données sont presque systématiquement communiquées).

⁶⁰ [Titre II, Chapitre IV Code de l'environnement](#)

Fiche 5 : Tableau comparatif des licences CC4.0 et CC0

	Attribution CC BY	Attribution - Pas de Modification CC BY-ND	Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions CC BY-NC-SA	Attribution - Partage dans les Mêmes Conditions CC BY-SA	Attribution - Pas d'Utilisation Commerciale CC BY-NC	Attribution - Pas d'Utilisation Commerciale - Pas de Modification CC BY-NC-ND	Domaine public - Universel CC0
							
Objet	<ul style="list-style-type: none"> • Propriété littéraire et artistique • Droit sui generis des bases de données : « right to extract, reuse, reproduce, and Share all or a substantial portion of the contents of the database » 						Tous les droits sauf : <ul style="list-style-type: none"> • Marque • Brevet • Interdiction légale.
Durée	Identique à la durée des droits de propriété intellectuelle sauf suppression ou modification de la licence. Les suppressions/modifications n'ont aucune incidence sur les versions précédentes déjà en cours d'exécution.						Illimitée.
Périmètre	Non-exclusif, monde entier.						
Gratuité	Oui						
Utilisation commerciale	Oui	Oui	Non	Oui	Non	Non	Oui
Modifications	Oui	Non	Oui	Oui	Oui	Non	Oui
Droits accordés	Reproduire et partager le contenu sous licence Produire,	Reproduire et partager le contenu sous licence Produire, reproduire sans partager du contenu	Reproduire et partager le contenu sous licence dans un but non commercial Produire, reproduire et	Reproduire et partager le contenu sous licence Produire, reproduire et partager du contenu	Reproduire et partager le contenu sous licence dans un but non commercial	Reproduire et partager le contenu sous licence dans un but non commercial	Reproduire et partager le contenu sous licence Produire, reproduire et partager du contenu

	reproduire et partager du contenu adapté Modifications techniques si nécessaire	adapté Modifications techniques si nécessaire	partager du contenu adapté dans un but non commercial Modifications techniques si nécessaire	adapté Modifications techniques si nécessaire	Produire, reproduire et partager du contenu adapté dans un but non commercial Modifications techniques si nécessaire	Produire, reproduire sans partager du contenu adapté dans un but non commercial Modifications techniques si nécessaire	adapté Modifications techniques si nécessaire
Droits moraux à respecter	<ul style="list-style-type: none"> • Paternité • Intégrité 						Aucun. Attention : en droit français le droit à la paternité s'applique même au-delà de la licence = obligation de citer le nom de l'auteur
Exceptions à la licence	Pas de mentions particulières. Mais quoi qu'il en soit : <ul style="list-style-type: none"> ❖ Europe = impossibilité d'interdire les extractions de la base de données : <ul style="list-style-type: none"> • à des fins scientifiques ou d'illustration pour l'enseignement ; • lorsque cela est requis par la loi ; • qualitativement ou quantitativement non substantielles de la base. ❖ Hors Europe = impossibilité d'interdire l'utilisation de la base si c'est pour un usage loyal ou « fair use » 						Aucune.
Autorisation de création dérivée	Oui	Oui mais partage interdit	Oui	Oui	Oui	Oui mais partage interdit	Oui
Partage de la création dérivée dans les mêmes conditions ou « share-alike »	Non	Non car partage interdit	Oui	Oui	Non	Non car partage interdit	Non
Conférer des	Non. L'auteur de la base initiale ne peut être tenu au-delà de la licence. Pour conférer des droits à un tiers, ce tiers doit accepter la licence ou contacter						Non précisé sauf cas

droits à un tiers sur la base initiale	l'auteur.	où la renonciation n'est pas valable = interdiction
Garanties	Aucune garantie sur le contenu faisant l'objet de la licence.	
Exclusion de responsabilité	Oui. Pas conforme au droit français. Exemple : exclusion de responsabilité du préjudice direct impossible en France.	Oui

Fiche 6 : Tableau comparatif des licences gratuites ODBL et Etalabab

	ODBL 1.0	Etalab (au 01/04/2015)
		Attention : cette licence ne peut être utilisée que s'il est possible de garantir que le contenu ne contient pas de droits de propriété intellectuelle appartenant à des tiers. Cela implique de contrôler chaque élément faisant objet de la licence, ce qui en fait une garantie très difficile à donner.
Objet	Base de données (= structure + données, mais pas les données isolément)	Informations (= la donnée)
Définitions	Nombreuses définitions	Peu de définitions
Durée	Durée des droits sur la base (= 15 ans sur le contenu)	Durée illimitée*
Périmètre	Non-exclusif, monde entier	Non-exclusif, monde entier
Gratuité	Oui	Oui
Utilisation commerciale	Oui	Oui
Droits d'utilisation accordés	<ul style="list-style-type: none"> • Extraire • Copier • Diffuser • Réutiliser • Créer des bases dérivées 	
	<ul style="list-style-type: none"> • Créer des bases collaboratives • Modifier pour des raisons techniques 	<ul style="list-style-type: none"> • Modifier
Sauf accord de l'administration, la réutilisation des informations publiques est soumise à la condition que ces dernières ne soient pas altérées, que leur sens ne soit pas dénaturé et que leurs sources et la date de leur dernière mise à jour soient mentionnées (article 12 de la loi n° 78-753 du 17 juillet 1978).		
Droits moraux à respecter	<ul style="list-style-type: none"> • Paternité • Respect de l'honneur de l'auteur • Respect de l'intégrité de l'œuvre, divulgation et repentir 	<ul style="list-style-type: none"> • Paternité (nom de l'auteur sur l'œuvre)

Exceptions à la licence	<ul style="list-style-type: none"> Hors Europe: Impossibilité d'interdire l'utilisation de la base si c'est pour un "usage loyal" (fair use) Europe : Extractions à des fins scientifiques ou d'illustration pour l'enseignement ou lorsque cela est requis par la loi Partout : Impossibilité d'interdire les extractions qualitativement ou quantitativement non substantielles de la base 	Pas de mentions
Création dérivée	Obligation de communication au public de la base de données dérivée (effet contaminant)	Aucune obligation de communication au public de l'information dérivée
Partage de la création dérivée	« Share alike » : obligation de partager toute création dérivée sous cette licence ODBL ou licence équivalente). Sauf exceptions art. 4.5	Pas de mentions
Conférer des droits à un tiers sur la base initiale	Non autorisé. Le tiers doit contacter le producteur initial de la base de données.	Possibilité de transmettre l'information à un tiers, sous la licence de son choix.
Garanties	Aucune	Garantie que l'information ne contient de droits de propriété intellectuelle tiers.
Exclusion de responsabilité	Oui (pas forcément conforme au droit français)	Oui (conforme au droit français)
Conclusion	Choisir cette licence si l'on a une base de données et que l'on souhaite contrôler les redistributions de la base et les travaux dérivés	Choisir cette licence si l'on n'a pas besoin d'un suivi sur le devenir des données et que les données sont essentiellement distribuées en France (bien que la licence soit compatible avec une distribution à l'étranger). Cette licence ne reprend pas les droits spécifiques en matière de base de données.

*Sauf extinction des droits du concédant sur le contenu

Fiche 7 : Mise en place d'une licence par la voie du contrat électronique

*Pour en savoir plus
Quelles sont les mentions
obligatoires sur un site internet ?*

Au-delà de la licence ou du contrat électronique, tout site internet édité à titre professionnel doit d'abord contenir un certain nombre de mentions légales. Le manquement à l'une de ces obligations peut être sanctionné jusqu'à un an d'emprisonnement et 75 000 € d'amende pour les personnes physiques ou 375 000 € pour les personnes morales.

Une licence est un contrat par lequel le titulaire des droits d'auteur sur un logiciel, une œuvre audiovisuelle, une base de données, etc., définit avec son cocontractant les conditions dans lesquelles l'œuvre peut être utilisée, diffusée ou modifiée.

Dans le cas spécifique des licences, il s'agit le plus souvent de contrats d'adhésion, c'est-à-dire d'un contrat dont les termes sont imposés par une partie à l'autre. Les clauses contractuelles sont fixées à l'avance et aucune discussion n'est possible : les cocontractants sont alors libres d'adhérer ou non à la licence en acceptant les termes du contrat tels quels ou bien en ne les acceptant pas du tout.

Il doit être clairement indiqué que le contenu mis à disposition est soumis à une licence. Par ailleurs, cette licence ou un lien direct vers celle-ci doit être visible et très facilement accessible. *Exemple : apposition d'un lien vers l'adresse URL de la licence juste en dessous de la phrase indiquant que le contenu est soumis à une licence.*

L'indication de la soumission du contenu à une licence peut se faire à l'aide d'une formule telle que : « le contenu est mis à disposition selon les termes de la licence [nom et identification de la licence]. Toute utilisation du contenu autre que celle autorisée par cette licence est interdite »⁶¹.

La saisie d'une « case à cocher » et le clic sur un bouton d'acceptation sont recommandés pour pouvoir organiser la preuve de l'acceptation de la licence. En pratique l'indication de la soumission du contenu à une licence se fait par une citation du nom de la licence choisie et de son logo :

- sur la première page du contenu lorsqu'il s'agit d'un document fixe ou non modifiable. *Exemple : page de garde d'un fichier « .pdf » ;*

⁶¹ Il est préférable d'utiliser une licence telle que les licences Creative Commons reprenant par ailleurs en son sein une formule telle que : « l'exercice de tout droit sur l'œuvre mise à disposition emporte acceptation des termes de la licence ».

- sur chaque page lorsqu'il s'agit d'un document dynamique tel qu'un site web ou une base de données.

Attention !! ces recommandations sont valables pour les licences gratuites. Pour les licences payantes, le formalisme est beaucoup plus lourd et je dois prendre contact avec mon service juridique.